

Understanding Cisco Express Forwarding

Document ID: 47321

Introduction

Prerequisites

Requirements

Components Used

Conventions

Overview

CEF Operations

Updating the GRP's Routing Tables

Packet Forwarding for all Line Cards Except OC48 and QOC12

Packet Forwarding for OC48 and QOC12 Line Cards

Related Information

Introduction

This document explains what Cisco Express Forwarding is, and how it is implemented in the Cisco 12000 Series Internet Router.

Prerequisites

Requirements

There are no specific requirements for this document.

Components Used

This document is not restricted to specific software and hardware versions.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, make sure that you understand the potential impact of any command.

Conventions

For more information on document conventions, see the Cisco Technical Tips Conventions.

Overview

Cisco Express Forwarding (CEF) switching is a proprietary form of scalable switching intended to tackle the problems associated with demand caching. With CEF switching, the information which is conventionally stored in a route cache is split up over several data structures. The CEF code is able to maintain these data structures in the Gigabit Route Processor (GRP), and also in slave processors such as the line cards in the 12000 routers. The data structures that provide optimized lookup for efficient packet forwarding include:

- The Forwarding Information Base (FIB) table – CEF uses a FIB to make IP destination prefix-based switching decisions. The FIB is conceptually similar to a routing table or information base. It maintains a mirror image of the forwarding information contained in the IP routing table. When

routing or topology changes occur in the network, the IP routing table is updated, and these changes are reflected in the FIB. The FIB maintains next-hop address information based on the information in the IP routing table.

Because there is a one-to-one correlation between FIB entries and routing table entries, the FIB contains all known routes and eliminates the need for route cache maintenance that is associated with switching paths such as fast switching and optimum switching.

- Adjacency table – Nodes in the network are said to be adjacent if they can reach each other with a single hop across a link layer. In addition to the FIB, CEF uses adjacency tables to prepend Layer 2 addressing information. The adjacency table maintains Layer 2 next-hop addresses for all FIB entries.

CEF can be enabled in one of two modes:

- Central CEF mode – When CEF mode is enabled, the CEF FIB and adjacency tables reside on the route processor, and the route processor performs the express forwarding. You can use CEF mode when line cards are not available for CEF switching, or when you need to use features not compatible with distributed CEF switching.
- Distributed CEF (dCEF) mode – When dCEF is enabled, line cards maintain identical copies of the FIB and adjacency tables. The line cards can perform the express forwarding by themselves, relieving the main processor – Gigabit Route Processor (GRP) – of involvement in the switching operation. This is the only switching method available on the Cisco 12000 Series Router.

dCEF uses an Inter-Process Communication (IPC) mechanism to ensure synchronization of FIBs and adjacency tables on the route processor and line cards.

For more information about CEF switching, see Cisco Express Forwarding (CEF) White Paper.

CEF Operations

Updating the GRP's Routing Tables

Figure 1 illustrates the process by which a routing update packet is sent to the Gigabit Route Processor (GRP) and the resulting forwarding update messages are sent to FIB tables on the line cards.

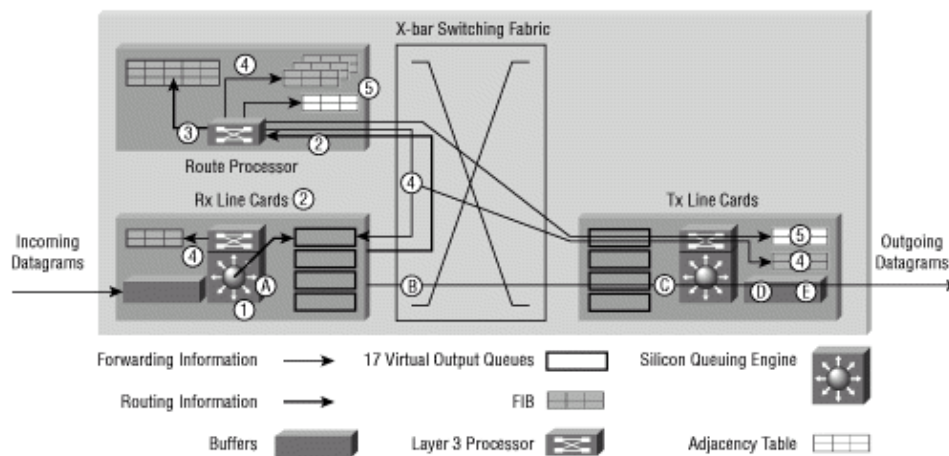
For clarity, the numbering of the following paragraphs corresponds to the numbering in Figure 1. The following process occurs during route table initialization, or any time the network topology changes (when routes are added, removed, or changed). The process shown in Figure 1 involves five main steps:

1. An IP datagram is placed into the input buffers on the receiving line card (ingress line card), and the L2/L3 forwarding engine accesses the Layer 2 and Layer 3 information in the packet and sends it to the forwarding processor. The forwarding processor determines that the packet contains routing information. The forwarding processor sends the pointer to the GRP virtual output queue (VOQ) indicating that the packet in buffer memory has to be sent to the GRP.
2. The line card issues a request to the clock and scheduler card (CSC). The scheduler card issues a grant, and the packet is sent across the switching fabric to the GRP.
3. The GRP processes the routing information. The R5000 (processor) on the GRP updates the network routing table. Depending on the routing information in the packet, the Layer 3 processor might have to flood link-state information to adjacent routers (if the internal routing protocol is Open Shortest Path First – OSPF). The processor generates the IP packets that carry the link-state information and the internal update for the FIB tables. Additionally, the GRP calculates all of the recursive routes that occur when support is provided for both an interior protocol and external gateway protocols (for example, Border Gateway Protocol – BGP).

The precalculated recursive route information is sent to the FIBs on each line card. This significantly speeds up the forwarding process, because the layer 3 processor on the line card can focus on forwarding the packet, not on calculating the recursive route.

4. The GRP sends out internal updates to FIB tables on all line cards, including those located on the GRP. The FIB updates to the line cards are monitored and throttled as needed. The GRP has a copy of each line card's FIB table, so if a new line card is inserted into the chassis, the GRP downloads the latest forwarding information to the new card, once that card becomes active.
5. The GRP is notified, from the line cards, whenever a new neighbor router is connected to the 12000 router. The processor on the line card sends a packet to the GRP containing the new layer 2 information (typically Point-to-Point Protocol (PPP) header information). The GRP uses this layer 2 information to update the adjacency table located on the GRP and on the line cards. Each line card adds this layer 2 information to each packet as the packet is sent from the 12000 router. A copy of the adjacency table is maintained on the GRP for initialization purposes.

Figure 1: Path Determination and Layer 3 Switching Diagram



Packet Forwarding for all Line Cards Except OC48 and QOC12

Once the line cards have enough forwarding information to determine the path through the switching fabric (for instance, the destination of the next hop), the 12000 router is ready to forward packets. The following steps outline the simple and fast forwarding technique used by the 12000 router (see Figure 1). For clarity, the lettering of the paragraphs corresponds to the lettering in Figure 1.

- **A.** An IP datagram is placed into the input buffers on the receiving line card (Rx line card), and the L2/L3 forwarding engine accesses the Layer 2 and Layer 3 information in the packet and sends it to the forwarding processor. The forwarding processor determines that the packet contains data and is not a routing update. Based on the Layer 2 and Layer 3 information in the FIB table, the forwarding processor sends the pointer to the appropriate line card's VOQ indicating that the packet in buffer memory is to be sent to that line card.
- **B.** The line card's scheduler issues a request to the scheduler. The scheduler issues a grant, and the packet is sent from the buffer memory across the switching fabric to the line card (Tx line card).
- **C.** The Tx line card buffers the incoming packets.
- **D.** The Layer 3 processor and associated application-specific integrated circuits (ASICs) on the Tx line card attach the Layer 2 information (a PPP address) to each transmitted packet. The packet is duplicated for each port on the line card (if needed).
- **E.** The Tx line card transmitters send the packet across the fiber interface.

The advantage of this simple forwarding process is that most data transmission tasks can be done in ASICs, allowing the 12000 to operate at gigabit rates. Also, data packets are never sent to the GRP.

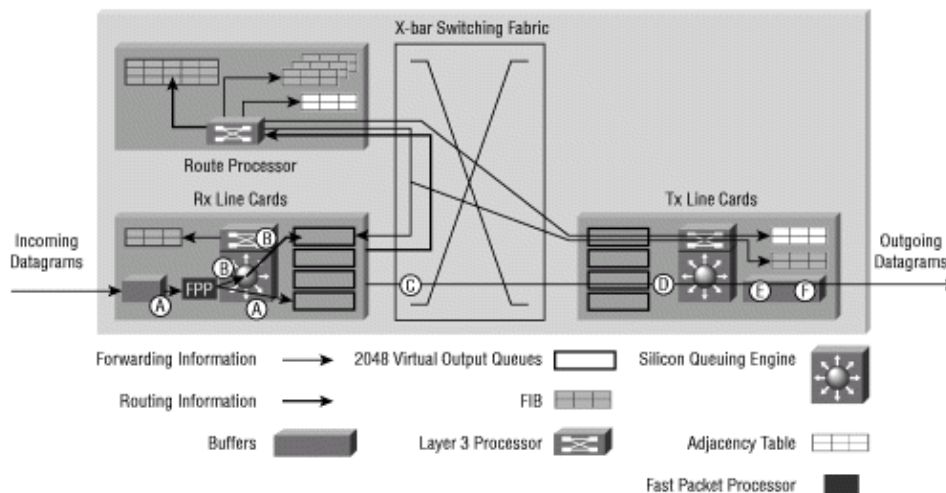
Packet Forwarding for OC48 and QOC12 Line Cards

When the line cards have enough forwarding information to determine the path through the switching fabric (for example, the destination of the next hop), the 12000 router is ready to forward packets. The following steps make up the simple and hyper-fast forwarding technique used by the 12000 (see Figure 2). For clarity, the lettering of the paragraphs corresponds to the lettering in Figure 2.

- **A.** An IP datagram (not a routing update, Internet Control Message Protocol (ICMP), and IP packets with options) is received into the line card and goes through layer 2 processing. Based on the layer 2 and layer 3 information in the local FIB table, the Fast Packet Processor determines the destination of the packet and modifies the packet header. Based on the destination, the packet is then placed in the appropriate line card's VOQ.
- **B.** In the rare case where the Fast Packet Processor cannot properly forward the packet, the packet is processed by the forwarding processor. The forwarding processor, based on the layer 2 and layer 3 information its local FIB table, sends the pointer to the appropriate line card's VOQ, indicating that the packet in buffer memory is to be sent to that line card.
- **C.** Once the packet is in the appropriate VOQ, the line card's scheduler issues a request to the scheduler. The scheduler issues a grant, and the packet is sent from the buffer memory across the switching fabric to the line card (Tx line card).
- **D.** The Tx line card buffers the incoming packets.
- **E.** The Layer 3 processor and associated ASICs on the Tx line card attach the Layer 2 information (a PPP address) to each transmitted packet. The packet is duplicated for each port on the line card (if needed).
- **F.** The Tx line card transmitters send the packet across the fiber interface.

The advantage of the new forwarding process is that it optimizes the card specifically for faster speeds, such as the OC48/STM16.

Figure 2: Packet Switching for Faster Line Cards



Related Information

- [Cisco 12000 Series Internet Router Architecture – Chassis](#)
- [Cisco 12000 Series Internet Router Architecture – Switch Fabric](#)
- [Cisco 12000 Series Internet Router Architecture – Route Processor](#)
- [Cisco 12000 Series Internet Router Architecture – Line Card Design](#)
- [Cisco 12000 Series Internet Router Architecture – Memory Details](#)

- **Cisco 12000 Series Internet Router Architecture – Maintenance Bus, Power Supplies and Blowers, and Alarm Cards**
 - **Cisco 12000 Series Internet Router Architecture – Software Overview**
 - **Cisco 12000 Series Internet Router Architecture – Packet Switching**
 - **Technical Support – Cisco Systems**
-

[Contacts & Feedback](#) | [Help](#) | [Site Map](#)

© 2008 – 2009 Cisco Systems, Inc. All rights reserved. [Terms & Conditions](#) | [Privacy Statement](#) | [Cookie Policy](#) | [Trademarks of Cisco Systems, Inc.](#)

Updated: Jan 17, 2006

Document ID: 47321
